

A Generalized Bidiagonal-Tikhonov Method Applied To Differential Phase Contrast Tomography

October 13, 2015

NICK SCHENKELS

Department of Mathematics and Computer Science,
University of Antwerp, Belgium
Middelheimlaan 1, M.G.328, 2020, Antwerpen, België

WIM VAN AARLE

Department of Physics, Vision Lab
University of Antwerp, Belgium
Universiteitsplein 1, D.N.114, 2610, Wilrijk, België

JAN SIJBERS

Department of Physics, Vision Lab
University of Antwerp, Belgium
Universiteitsplein 1, D.N.113, 2610, Wilrijk, België

WIM VANROOSE

Department of Mathematics and Computer Science,
University of Antwerp, Belgium
Middelheimlaan 1, M.G.210, 2020, Antwerpen, België

Abstract

Phase contrast tomography is an alternative to classic absorption contrast tomography that leads to higher contrast reconstructions in many applications. We review how phase contrast data can be acquired by using a combination of phase and absorption gratings. Using algebraic reconstruction techniques the object can be reconstructed from the measured data. In order to solve the resulting linear system we propose the Generalized Bidiagonal Tikhonov (GBiT) method, an adaptation of the generalized Arnoldi-Tikhonov method that uses the bidiagonal decomposition of the matrix instead of the Arnoldi decomposition. We also study the effect of the finite difference operator in the model by examining the reconstructions with either a forward difference or a central difference approximation. We validate our conclusions with simulated and experimental data.

1 Introduction

Since their discovery by Wilhelm Röntgen in 1895, x-rays have become an invaluable tool in medical imaging, material testing, security screenings, etc. In classic x-ray

computed tomography (CT), the absorption of the x-rays by the material is imaged. However, very different materials can have similar absorption effects, making them hard to differentiate in the reconstructed image. An alternative tomographic technique is phase contrast computed tomography (PC-CT). This method is based on the phase shift that occurs in the wave front when the x-rays pass through the object of interest. This shift is more sensitive to small changes in the material – especially for soft tissues or samples with low atomic numbers – this technique can be used to achieve a higher contrast in the reconstructions.

The reconstruction methods for absorption contrast CT (AB-CT) are well established in literature and can be divided in two groups. First there are analytical techniques, such as filtered back projection (FBP), which are based on the Radon transform and its relation to the Fourier transform as expressed by the Fourier Slice theorem [26]. The second type are algebraic reconstruction methods, which are based on the discretization of the Radon transform $R \in \mathbb{R}^{m \times n}$ and solve a linear system

$$Rx = b, \quad (1)$$

using some direct or iterative method. Here, $x \in \mathbb{R}^n$ represents a rasterized version of the attenuation coefficients of the object and $b \in \mathbb{R}^m$ the measured x-ray intensity data. In practice, b is subject to noise and measurement errors, which can result in various types of reconstruction artefacts. Moreover, constraints on radiation dose, scanning time and scanning geometry can make (1) an underdetermined system, leading to non-unique solutions. Various regularization methods have been developed in order to find a good approximation of the image x . These methods typically include the exploitation of prior knowledge about the scanned object. One widely used approach is Tikhonov regularization, which solves:

$$x_\lambda = \arg \min_{x \in \mathcal{K}} \left\{ \|Rx - b\|^2 + \lambda \|L(x - x_0)\|^2 \right\},$$

where $\|\cdot\|$ denotes the 2-norm and \mathcal{K} is a subspace of the data-space \mathbb{R}^n . Prior knowledge of the solution is incorporated via the initial guess x_0 and the operator L , which can be used to place constraints on the solution x_λ . The choice of the regularization parameter λ plays a key role here, since it determines the quality of the reconstruction. We come back to this in [section 3](#).

Differential phase contrast tomography

In traditional absorption contrast CT (AC-CT), the intensity of the x-rays is measured immediately after they pass through the object of interest. However, this doesn't provide any information about the occurred phase shift, information which can lead to a significantly enhanced contrast between two materials with similar absorption coefficients. Various different approaches have been developed in order to acquire data that can be linked to the phase shift: set-ups with a crystal interferometer [2, 23], propagation based imaging [4, 21, 33], analyser based imaging [18, 19] or differential phase contrast imaging [22, 32].

Differential phase contrast CT (DPC-CT) is based on the small angular deviations α that appear in the wave front after going through the scanned object, which in turn can be linked to the derivative of the phase shift Φ . Recovery of these refraction angles based on the intensity measurements can be done using a reverse projection method [36], a Fourier based algorithm [34] or a phase stepping approach.

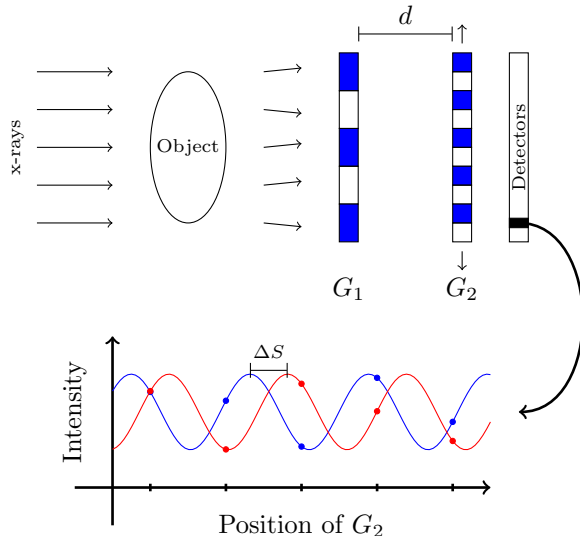


Figure 1: Overview of the DPC set-up. The scanned object causes a phase shift in the wave front. The angular deviations this causes in the x-rays are proportional to the differential phase shift. By using a phase grating (G_1) and an absorption grating (G_2) with a phase-stepping procedure, it is possible to measure these small angles.

In the latter, two gratings are placed between the object and the detector array, see Fig. 1. The first grating G_1 is a phase grating that, due to the Talbot effect and the Fresnel diffraction principle, causes an interference pattern at the Talbot distance d [3]. In order to accurately measure this interference pattern, a second absorption grating G_2 is used. For different positions of G_2 , measurements of the intensity of the outgoing x-rays are taken with and without an object present in the detector. By comparing the two intensity curves (position of G_2 versus the intensity measurement), the shift ΔS caused by the object can be measured. Since ΔS is related to α , the refraction angles can be recovered. For a more detailed overview of this set-up and the phase stepping approach we refer to [24, 32].

The reconstruction methods for DPC-CT are closely related to those for AB-CT. Analytically, the presence of a derivative of the Radon transformation in the FBP theorem results in a special complex valued filter. Algebraically on the other hand, the model can, for example, be represented using a discrete version for the derivative of the Radon transform [27] or by adding a finite difference operator D as proposed in [9, 10, 11]:

$$DRx = b.$$

Here, the resulting linear system is solved using a Kaczmarz method or algebraic reconstruction technique (ART).

In this paper we follow a similar algebraic approach, but we propose to use an adaptation of the generalized Arnoldi-Tikhonov method to solve the linear system. This is a Krylov subspace method that uses Tikhonov regularization and an secant update step based on the discrepancy principle in order to find a suitable value for the regularization parameter λ . This method was first introduced in [12, 13] where it was derived for square matrices. As we will illustrate further on, square matri-

ces rarely appear in tomographic reconstructions – absorption or phase contrast – but the method can be adjusted using a bidiagonal decomposition of the matrix instead of the Arnoldi algorithm. Note that this generalization is independent of the tomographic setting we use as a test case. We will also take a closer look at how the choice of finite difference operator D effects the reconstruction. We do this by comparing the forward difference scheme used in [9, 10, 11] with a central difference approximation and study the effects they both have on the reconstruction.

Outline

The outline of this paper is as follows. In [section 2](#) we give a brief review of absorption and phase contrast tomography and describe how they can be solved analytically and algebraically. In [section 3](#) we discuss the Arnoldi-Tikhonov method and explain how it can be adjusted for non-square matrices. In [section 4](#) we examine how the finite difference operator affects the solution by looking at a simple test problem. Finally, in [section 5](#), we look at some numerical experiments with simulated and experimental data.

2 Review of the forward and inverse problem

The propagation of a wave $u(\mathbf{x})$ through a medium with no inside source is described by the homogeneous Helmholtz equation

$$(\Delta + k(\mathbf{x})) u(\mathbf{x}) = 0$$

with $\mathbf{x} \in \mathbb{R}^2$ and $k(\mathbf{x}) \in \mathbb{C}$ the position dependent wave number. If θ is the direction of the incoming plane wave $u_{in}(\mathbf{x}) = E_0 e^{ik_0 \mathbf{x} \cdot \boldsymbol{\theta}}$, with $\boldsymbol{\theta} = (\cos(\theta), \sin(\theta))^T$ and amplitude E_0 , then $k(\mathbf{x}) = k_0 n(\mathbf{x})$. Here,

$$n(\mathbf{x}) = 1 - \delta(\mathbf{x}) + i\beta(\mathbf{x}) \in \mathbb{C}$$

is the refractive index of the material and $\beta, \delta : \mathbb{R}^2 \rightarrow \mathbb{R}$. Assuming that the wave number is constant, i.e. $k(\mathbf{x}) = k = k_0(1 - \delta + i\beta)$, then we can write the total wave as

$$\begin{aligned} u(\mathbf{x}) &= E_0 e^{ik\mathbf{x} \cdot \boldsymbol{\theta}} \\ &= \left(E_0 e^{ik_0 \mathbf{x} \cdot \boldsymbol{\theta}}\right) \left(e^{-ik_0 \delta \mathbf{x} \cdot \boldsymbol{\theta}}\right) \left(e^{-k_0 \beta \mathbf{x} \cdot \boldsymbol{\theta}}\right) \\ &= \left(\text{incoming wave}\right) \left(\text{phase shift}\right) \left(\text{amplitude change}\right). \end{aligned} \tag{2}$$

This equation illustrates the effect of the medium on the incoming wave: the imaginary part β of the refractive index causes a change in the amplitude of the wave, whereas the real part δ leads to a phase shift in the wave front. From (2) it follows that the intensity of the beam at \mathbf{x} is given by

$$\begin{aligned} I(\mathbf{x}) &= \left(\text{amplitude of the wave}\right)^2 \\ &= \left(E_0 e^{-k_0 \beta \mathbf{x} \cdot \boldsymbol{\theta}}\right)^2 \\ &= I_0 e^{-2k_0 \beta \mathbf{x} \cdot \boldsymbol{\theta}}, \end{aligned}$$

with $I_0 = E_0^2$ the intensity of the incoming wave. If the beam travels from \mathbf{x} to $\mathbf{x} + \Delta\mathbf{x}$ the change in intensity can be approximated by

$$\begin{aligned} I(\mathbf{x} + \Delta\mathbf{x}) &\approx I(\mathbf{x}) + I'(\mathbf{x})\Delta\mathbf{x} \cdot \boldsymbol{\theta} \\ &= I(\mathbf{x}) - 2k_0\beta I(\mathbf{x})\Delta\mathbf{x} \cdot \boldsymbol{\theta} \\ \Rightarrow dI(\mathbf{x}) &= -2k_0\beta I(\mathbf{x})d\mathbf{x} \cdot \boldsymbol{\theta}. \end{aligned} \quad (3)$$

Similarly, if $\Phi(\mathbf{x})$ is the phase shift of the wave, then the phase shift in the wave propagating from \mathbf{x} to $\mathbf{x} + \Delta\mathbf{x}$ is given by

$$\begin{aligned} \Phi(\mathbf{x} + \Delta\mathbf{x}) - \Phi(\mathbf{x}) &= k_0\delta\Delta\mathbf{x} \\ \Rightarrow d\Phi(\mathbf{x}) &= k_0\delta d\mathbf{x} \cdot \boldsymbol{\theta}. \end{aligned} \quad (4)$$

Note that this is the phase shift relative to the incoming wave. To find the intensity and total phase shift at the detector for a non-constant wave number $k(\mathbf{x})$, we have to integrate (3) and (4) along the entire path of the beam $L(\theta, t) = \{\mathbf{x} \in \mathbb{R}^2 \mid \mathbf{x} \cdot \boldsymbol{\theta} = t\}$, which yields:

$$\begin{cases} I(\theta, t) = I_0 e^{-2k_0 \int_{L(\theta, t)} \beta(\mathbf{x}) d\mathbf{x}} \\ \Phi(\theta, t) = \int_{L(\theta, t)} k_0 \delta(\mathbf{x}) d\mathbf{x} \end{cases}. \quad (5)$$

It should be noted that this is a very crude, linear approximation of the Helmholtz equation we started from. However, (5) is much easier to solve and can still be used to reconstruct high quality x-ray images.

2.1 Analytical Reconstruction

Analytically, AB-CT is based on the filtered back projection theorem, which states that [26]:

$$\beta(\mathbf{x}) = \mathcal{B}(\mathcal{F}^{-1}|r|\mathcal{F}\mathcal{R}\beta)(\mathbf{x}). \quad (6)$$

Here, \mathcal{F} is the Fourier transform, \mathcal{R} and \mathcal{B} the Radon transformation and its inverse and r the radius of \mathbf{x} in polar coordinates. Since it follows from (5) that

$$\mathcal{R}\beta(\theta, t) = \int_{L(\theta, t)} \beta(\mathbf{x}) d\mathbf{x} = -\frac{1}{2k_0} \ln \frac{I(\theta, t)}{I_0}$$

it is possible to reconstruct β via (6) using the measurements of $I(\theta, t)$.

Differential phase contrast tomography, however, is based on the small angular deviations $\alpha(\theta, t)$ that occur in the wave front. By applying a phase stepping technique as described in section 1, the shift in the intensity measurements caused by the object can be determined. Denoting this shift as $\Delta S(\theta, t)$, they can be linked to the deviation angles in the x-rays by

$$\Delta S(\theta, t) = d\alpha(\theta, t).$$

Since these angles are very small, they can be approximated by

$$\alpha(\theta, t) = \frac{1}{k_0} \frac{\partial \Phi(\theta, t)}{\partial t}.$$

Substituting (5) for $\Phi(\theta, t)$, we find that:

$$\alpha(\theta, t) = \frac{\partial \int_{L(\theta, t)} \delta(x, y) ds}{\partial t} = \frac{\partial}{\partial t} \mathcal{R} \delta(\theta, t). \quad (7)$$

This differs from AB-CT because here we have the derivative of the Radon transform. However, the Fourier transform property for derivations states that

$$\mathcal{F} \left(\frac{df}{dx} \right) (r) = 2\pi r i \mathcal{F} f(r),$$

so it follows that

$$\mathcal{F} \alpha(\theta, r) = \mathcal{F} \left(\frac{\partial}{\partial t} \mathcal{R} \delta \right) (\theta, r) = 2\pi r i \mathcal{F} \mathcal{R} \delta(\theta, r).$$

This means that δ can be reconstructed similarly to (6), but with a different complex filter and measurements of $\alpha(\theta, t)$:

$$\delta(\mathbf{x}) = \mathcal{B} \left(\mathcal{F}^{-1} \frac{|r|}{2\pi r i} \mathcal{F} \alpha \right) (\mathbf{x}). \quad (8)$$

2.2 Algebraic Reconstruction

In both absorption and differential phase contrast tomography the measurements are related to the Radon transformation, a set of line integrals over the domain. Using a parallel beam projection model, see Fig. 2 these can be discretized in the following way:

- Assume that there are k detector elements and that projections are made under l different angles $\theta_1, \dots, \theta_l \in [0, \pi[$. This means there are $m = kl$ different line integrals or measurements, which can be written as a vector $b = (b_i)_{i=1}^m \in \mathbb{R}^m$. Each component b_i will be the measurement corresponding to a specific detector element and projection angle.
- The domain is considered to be a grid of n_x by n_y pixels in the x and y direction, which are written column wise as a vector $x = (x_j)_{j=1}^n \in \mathbb{R}^n$ with $n = n_x n_y$.
- For a measurement b_i , every pixel x_j is given a weight $R = (r_{ij})_{ij} \in \mathbb{R}^{m \times n}$ such that

$$b_i = \sum_{j=1}^n r_{ij} x_j \approx \text{the corresponding line integral.}$$

The weights r_{ij} can be chosen in different ways, see [20], but here we take them to be the length of the line segment of projection b_i going through pixel x_j , see Fig. 2. For our numerical experiments in sections 4 and 5 we generate the matrix R on the fly using the ASTRA toolbox [30, 35].

The resulting linear system $Rx = b$, for $x \in \mathbb{R}^n, b \in \mathbb{R}^m$ and $R \in \mathbb{R}^{m \times n}$, is a discrete version of the Radon transformation and can be used to solve the absorption based CT problem. However, in DPC-CT we use measurements of the derivative of

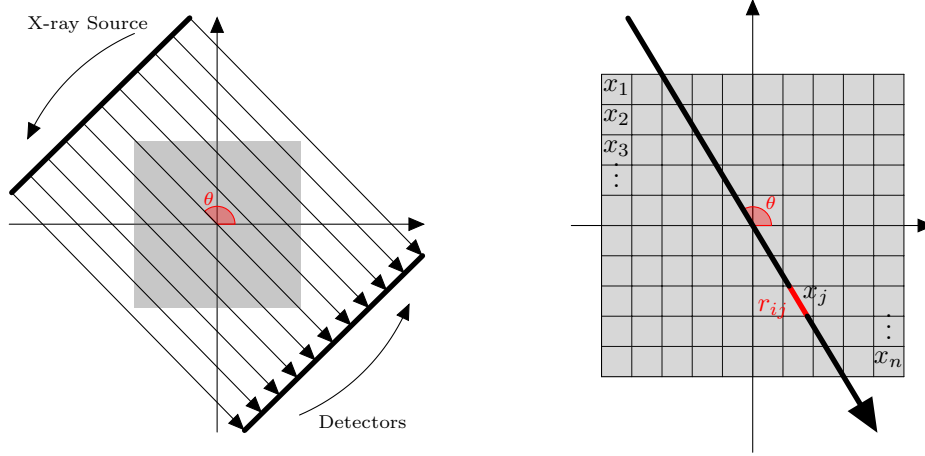


Figure 2: Overview of parallel beam projection & definition of the weights.

the Radon transform. This results in a new linear system $DRx = b$, where D is a finite difference matrix.

Two natural choices for D are the forward or the central finite difference approximation. If we set

$$q(\theta, t) = \mathcal{R} \delta(\theta, t) = \int_{L(\theta, t)} \delta(\mathbf{x}) d\mathbf{x},$$

and take h to be the distance between the detector cells, the forward and central difference scheme are respectively given by:

$$\frac{\partial q(\theta, t)}{\partial t} = \frac{q(\theta, t+h) - q(\theta, t)}{h} + \mathcal{O}(h)$$

or

$$\frac{\partial q(\theta, t)}{\partial t} = \frac{q(\theta, t+h) - q(\theta, t-h)}{2h} + \mathcal{O}(h^2).$$

So the discrete values of α are approximated by one of the following schemes:

$$\alpha(\theta_i, t_j) = \left. \frac{\partial q(\theta_i, t)}{\partial t} \right|_{t=t_j} = \frac{q(\theta_i, t_{j+1}) - q(\theta_i, t_j)}{h}$$

or

$$\alpha(\theta_i, t_j) = \left. \frac{\partial q(\theta_i, t)}{\partial t} \right|_{t=t_j} = \frac{q(\theta_i, t_{j+1}) - q(\theta_i, t_{j-1})}{2h}.$$

The matrices D corresponding with these approximations are given below. Note that both are constructed using a Kronecker product because the derivative in (7) is taken with respect to the detector space and independently of the projection angle. In [section 4](#) we will take a closer look at how each matrix affects the reconstruction.

- Forward difference scheme:

$$D_f = I_l \otimes \underbrace{\begin{pmatrix} -1 & 1 & & & \\ & -1 & 1 & & \\ & & -1 & 1 & \\ & & & \ddots & \ddots \\ & & & & -1 & 1 \\ & & & & & -1 \end{pmatrix}}_{k \times k} \in \mathbb{R}^{m \times m}.$$

- Central difference scheme:

$$D_c = I_l \otimes \frac{1}{2} \underbrace{\begin{pmatrix} 0 & 1 & & & \\ -1 & 0 & 1 & & \\ & -1 & 0 & 1 & \\ & & \ddots & \ddots & \ddots \\ & & & -1 & 0 & 1 \\ & & & & -1 & 0 \end{pmatrix}}_{k \times k} \in \mathbb{R}^{m \times m}.$$

3 Iterative reconstruction with Krylov subspaces

One of the problems when dealing with inverse problems is that the forward problem still contains some model and approximation errors and that the right hand side b typically contains small measurement errors, i.e.

$$Ax = b = \tilde{b} + e$$

where $\tilde{b} = A\tilde{x}$ for the exact original image \tilde{x} . When using an iterative method to solve these problems, the first few iterations usually seem to converge very nicely to \tilde{x} . However, as the iterations continue, the noise starts to have an increasing effect on the approximations x_k and the error with respect to \tilde{x} increases again. This effect is called semi-convergence [17]. Therefore, some form of regularization method has to be performed in order to acquire a good approximation for \tilde{x} . An often used method for large scale linear problems is Tikhonov regularization, which instead of solving $Ax = b$, solves the problem

$$x_\lambda = \arg \min_{x \in \mathcal{K}} \left\{ \|Ax - b\|^2 + \lambda \|L(x - x_0)\|^2 \right\} \quad (9)$$

for some subspace $\mathcal{K} \subseteq \mathbb{R}^n$. Here, we will take \mathcal{K} to be the Krylov subspace spanned by the matrix $A^T A \in \mathbb{R}^{m \times n}$ and $A^T r_0$ for the initial residue $r_0 = b - Ax_0 \in \mathbb{R}^m$. This method can be derived from the singular value decomposition of A [17]. Intuitively, the idea is to balance solving the original problem – by minimizing the fidelity term $\|Ax - b\|^2$ – and suppressing the noise – by minimizing the penalty term $\|L(x - x_0)\|^2$. This is done by choosing an appropriate value for the regularization parameter λ . If prior knowledge of the solution is available, the choice of a regularization matrix L that incorporates these conditions in the solution can greatly improve the convergence of the method. Various variations exist to this

standard formulation, e.g. replacing the penalizing term by $\|x\|_1^2$ or with the total variation operator $\text{TV}(x)$. However, the difficulty lies in finding an optimal value for λ . If it is chosen too small, the method will still exhibit semi-convergence, and, if it is chosen too large, the solution will tend to be over-smooth.

There exist a number of different methods to find a value for λ that is in some way suitable: the L-curve [6, 7], the L-ribbon [5], generalized cross validation [8, 15], the discrepancy principle [25], In [12, 13] a generalized Arnoldi-Tikhonov method is proposed to solve (9) that incorporates an iterative scheme based on the discrepancy principle to find an optimal value for λ . However, since this method is based on the Arnoldi decomposition of a matrix, it only works for square matrices A . This will not be true for general inverse problem, e.g. in tomographic reconstructions, where the matrix A is non square due to the number of detectors, projection angles and pixels (see section 2.2). We therefore suggest an adaptation of this algorithm that uses a bidiagonal decomposition of the matrix A instead of the Arnoldi decomposition. Since both decompositions are closely connected, the Tikhonov problem (9) can be solved in a similar way.

3.1 Generalized Bidiagonal-Tikhonov

In [16], it is shown that a matrix $A \in \mathbb{R}^{m \times n}$ can be decomposed as

$$A = U \begin{pmatrix} B \\ 0 \end{pmatrix} V^T, \quad (10)$$

with $U \in \mathbb{R}^{m \times m}$ and $V \in \mathbb{R}^{n \times n}$ orthonormal, $0 \in \mathbb{R}^{(m-n) \times n}$ and $B \in \mathbb{R}^{(n+1) \times n}$ an lower bidiagonal matrix. We now denote the columns of U and V respectively as $u_1, \dots, u_m \in \mathbb{R}^m$ and $v_1, \dots, v_n \in \mathbb{R}^n$ and the matrices of the first k columns as $U_k = (u_1, \dots, u_k)$, $V_k = (v_1, \dots, v_k)$ and $B_{k+1,k}$. If

$$B = \begin{pmatrix} \alpha_1 & & & \\ \beta_2 & \alpha_2 & & \\ & \beta_3 & \ddots & \\ & & \ddots & \alpha_n \\ & & & \beta_{n+1} \end{pmatrix} \in \mathbb{R}^{(n+1) \times n},$$

then (10) implies the following relations:

$$AV = U_n B \quad A^T U_n = V B^T.$$

Setting $\beta_1 v_0 = \alpha_{n+1} v_{n+1} = 0$, this means that for all $k = 1, 2, \dots, n$:

$$\begin{aligned} A^T u_k &= \beta_k v_{k-1} + \alpha_k v_k \\ Av_k &= \alpha_k u_k + \beta_{k+1} u_{k+1}. \end{aligned}$$

Starting with a given vector $u_1 \in \mathbb{R}^m$ and $\|u_1\| = 1$ we get the following recursion by Paige and Saunders called Bidiag1 [28, 29]:

$$\begin{cases} r_k = A^T u_k - \beta_k v_{k-1}, & \alpha_k = \|r_k\|, & v_k = \frac{1}{\alpha_k} r_k \\ p_k = Av_k - \alpha_k u_k, & \beta_{k+1} = \|p_k\|, & u_{k+1} = \frac{1}{\beta_{k+1}} p_k \end{cases} \quad (11)$$

This process stops when either $\alpha_k = 0$ or $\beta_k = 0$. The former can only happen when $\text{rank}(A^T) < m$ and the latter implies that

$$AV_{k-1} = U_{k-1}B_{k-1,k-1} \quad A^T U_{k-1} = V_{k-1}B_{k-1,k-1}^T,$$

meaning that the singular values of $B_{k,k-1}$ are a subset of the singular values of A [1]. After k steps of Bidiag1 with initial vector $u_1 = r_0 = b - Ax_0$, orthonormal matrices $V_k \in \mathbb{R}^{n \times k}$ and $U_{k+1} \in \mathbb{R}^{m \times (k+1)}$ and a lower bidiagonal matrix $B_{k+1,k} \in \mathbb{R}^{(k+1) \times k}$ have been constructed that satisfy

$$AV_k = U_{k+1}B_{k+1,k}.$$

Where the generalized Arnoldi-Tikhonov algorithm was intrinsically connected to the GMRES algorithm – which is based on the Arnoldi decomposition – the method we describe here will be linked to the LSQR algorithm – which is based on the bidiagonal decomposition. In each iteration, the current approximation is given by

$$x_{k,\lambda} = \arg \min_{x_k \in \mathcal{K}_k} \left\{ \|Ax_k - b\|^2 + \lambda \|L(x_k - x_0)\|^2 \right\},$$

where $x_{k,\lambda}$ lies in the Krylov subspace $\mathcal{K}_k = \mathcal{K}_k(A^T A, A^T r_0) = \text{span } V_k$. This means that the iterations can be written as $x_{k,\lambda} = x_0 + V_k y_{k,\lambda}$ for some $y_{k,\lambda} \in \mathbb{R}^n$. It's therefore equivalent in each iteration to solve

$$\begin{aligned} y_{k,\lambda} &= \arg \min_{y_k \in \mathbb{R}^k} \left\{ \|A(x_0 + V_k y_k) - b\|^2 + \lambda \|LV_k y_k\|^2 \right\} \\ &= \arg \min_{y_k \in \mathbb{R}^k} \left\{ \|AV_k y_k - r_0\|^2 + \lambda \|LV_k y_k\|^2 \right\} \\ &= \arg \min_{y_k \in \mathbb{R}^k} \left\{ \|U_{k+1}B_{k+1,k}y_k - r_0\|^2 + \lambda \|LV_k y_k\|^2 \right\} \\ &= \arg \min_{y_k \in \mathbb{R}^k} \left\{ \|B_{k+1,k}y_k - \|r_0\|e_1\|^2 + \lambda \|LV_k y_k\|^2 \right\} \\ &= \arg \min_{y_k \in \mathbb{R}^k} \left\| \begin{pmatrix} B_{k+1,k} \\ \sqrt{\lambda}LV_k \end{pmatrix} y_k - \begin{pmatrix} \|r_0\|e_1 \\ 0 \end{pmatrix} \right\|^2. \end{aligned} \quad (12)$$

Note that in the next to last equation we used the orthogonality of the columns of U_{k+1} and the fact that $u_1 = r_0/\|r_0\|$ and if L is the identity matrix, (12) can be written without the matrix V_k . Furthermore, if the regularization term $\|L(x - x_0)\|^2$ were dropped, this would simply be the LSQR algorithm.

3.2 Secant update step

We now describe how the parameter λ can be updated iteratively by using a approximation for the discrepancy in each iteration. This method was suggested in [12, 13] for the generalized Arnoldi-Tikhonov method. Due to the resemblance with the bidiagonal decomposition, analogous derivations can be made. The idea behind the discrepancy principle is to choose the regularization parameter λ such that

$$\|b - Ax_\lambda\| = \eta\varepsilon, \quad (13)$$

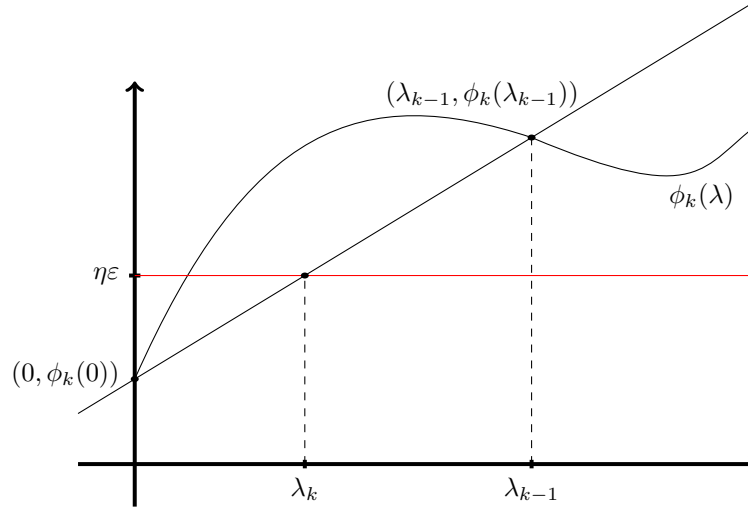


Figure 3: Overview of the secant update step.

with the error norm $\varepsilon = \|e\| = \|b - \tilde{b}\|$ and a small tolerance value $\eta \gtrsim 1$. We now define the discrepancy as the residual in an iteration as a function of λ

$$\phi_k(\lambda) = \|b - Ax_{k,\lambda}\|,$$

and say that the discrepancy principle is satisfied when $\phi_k(\lambda) \leq \eta\varepsilon$. In order to find a good approximation for λ in iteration k , we can perform one step of the secant method in order to approximate a solution for $\phi_k(\lambda) = \eta\varepsilon$. Assuming we have found a value λ_{k-1} in the previous iteration, this can be done using the initial points $(0, \phi_k(0))$ and $(\lambda_{k-1}, \phi_k(\lambda_{k-1}))$, see Fig. 3. Using these points, the update formula for λ in each iteration with the secant method is given by:

$$\lambda_k = \frac{\eta\varepsilon - \phi_k(0)}{\phi_k(\lambda_{k-1}) - \phi_k(0)} \lambda_{k-1}. \quad (14)$$

Therefore, in order to perform the secant update step, (12) has to be solved twice. Once with $\lambda = \lambda_{k-1}$, the current best estimate, and once with $\lambda = 0$. This means that $\phi_k(0)$ is simply the residual in the k -th iteration of the non-regularized LSQR method and can be found by solving

$$\min_{y \in \mathbb{R}^k} \|c - B_{k+1,k}y\|, \quad (15)$$

with $c = \|r_0\| e_1$. Because in the first few iterations the LSQR residual will be bigger than the tolerance value, i.e. $\phi_k(0) > \eta\varepsilon$, (14) would give a negative value for the regularization parameter, therefore (14) is replaced by

$$\lambda_k = \left| \frac{\eta\varepsilon - \phi_k(0)}{\phi_k(\lambda_{k-1}) - \phi_k(0)} \right| \lambda_{k-1} \quad (16)$$

This is also why in the first few iterations the value for λ typically becomes very large. Only after $\phi_k(0)$ becomes smaller than $\eta\varepsilon$, the value of the parameter will

improve. This is why the initial value of λ_0 has little to no effect on the final solution. In our numerical experiments we will therefore take $\lambda_0 = 1$. We solve two linear systems in each iteration, namely (12) and (15), but these are two sparse systems due to the bidiagonal structure of the matrix $B_{k+1,k}$ that is used in both systems. However, when the LSQR algorithm needs many iterations for its residual to go below the threshold value $\eta\varepsilon$, and hence for the regularization parameter to converge, these linear systems become larger and more compute time is required.

An overview of this method is given in Algorithm 1. Here, an extra condition is added to the stop criterion in order to ensure the convergence of λ_k . This is done by demanding that the discrepancy principle is satisfied a number of times before quitting.

Algorithm 1 Generalized Bidiagonal-Tikhonov (GBiT)

```

1: Input:  $A, b, x_0, \eta, \epsilon, \lambda_0, L$  and maxcounter
2: counter = 0
3: for  $k = 1, \dots, \text{maxIter}$  do
4:   1. Update the bidiagonal decomposition by (11) by calculating  $v_k$  and
       $u_{k+1}$  with reorthogonalization.
      2. Solve (15) and calculate  $\phi_k(0)$ .
      3. Solve (12) and calculate  $\phi_k(\lambda_{k-1})$ .
      4. Calculate  $\lambda_k$  by (16).
5:   if  $\phi_k(\lambda_{k-1}) < \eta\varepsilon$  then
6:     counter++
7:     if maxcounter < counter then
8:       break
9:     end if
10:  end if
11: end for
12:  $x_k = x_0 + V_k y_k$ 

```

3.3 Alternative update scheme

A downside of this iterative scheme for the regularization parameter, and of the discrepancy principle in general, is that a fairly good estimate of the noise norm ε is necessary. If it too small, the method will still exhibit semi-convergence, and if it is too big, the solution will be over-smoothed. This can be solved by changing the update formula (16) to

$$\lambda_k = \left| \frac{\eta\phi_{k-1}(0) - \phi_k(0)}{\phi_k(\lambda_{k-1}) - \phi_k(0)} \right| \lambda_{k-1}. \quad (17)$$

as was suggested in [14]. The motivation for this adjustment is that the non-regularized residual of LSQR tends to stabilize slightly below the noise norm, i.e. $\phi_{k-1}(0) \approx \varepsilon$. The only problem with (17) is that it may prevent the discrepancy principle from being satisfied. The reason for this is that when the approximation

for the regularization parameter starts to stabilize, the following will hold:

$$\begin{aligned}\lambda_k \approx \lambda_{k-1} &\Rightarrow \phi_k(\lambda_{k-1}) - \phi_k(0) \approx \eta\phi_{k-1}(0) - \phi_k(0) \\ &\Rightarrow \phi_k(\lambda_{k-1}) \approx \eta\phi_{k-1}(0)\end{aligned}$$

Therefore, we propose to take a slightly bigger tolerance value for the discrepancy principle than for the update step. For example, by taking its value 1% bigger, the condition on line 5 of [Algorithm 1](#) becomes

$$\phi_k(\lambda_{k-1}) < 1.01\eta\phi_{k-1}(0).$$

4 Finite difference approximation

In [section 2.2](#), we introduced the finite difference operators D_f and D_c . In this section, we will examine how these affect the reconstruction by studying a test problem. The following proposition summarizes some properties of the D_f and D_c which can easily be verified using induction and the properties of the kronecker product.

Proposition 1 (i) For $D_f \in \mathbb{R}^{m \times m}$ the following holds:

$$\det(D_f) = (-1)^m \quad \text{and} \quad \text{Null}(D_f) = \{0\}.$$

(ii) For $D_c^m \in \mathbb{R}^{m \times m}$ the following holds:

$$\det(D_c^m) = \begin{cases} \frac{1}{2^m} & m \text{ even} \\ 0 & m \text{ odd.} \end{cases}$$

and

$$\text{Null}(D_c^m) = \begin{cases} \{0\} & m \text{ even} \\ \langle (1, 0, 1, 0, \dots, 1, 0)^T \rangle & m \text{ odd} \end{cases}$$

Note that the inverse of D_f is given by the kronecker product of I_l and a backsubstitution matrix:

$$D_f^{-1} = I_l \otimes \underbrace{\begin{pmatrix} -1 & -1 & \dots & -1 \\ & -1 & \dots & -1 \\ & & \ddots & \vdots \\ & & & -1 \end{pmatrix}}_{k \times k} \in \mathbb{R}^{m \times m}.$$

4.1 Test problem

In order to study the effect of the finite difference operator on the reconstruction, we take the 256×256 Shepp-Logan phantom in MATLAB, see [Fig. 4a](#) and look at a single projection on 256 detectors (angle $\theta = \pi/2$). We take $x \in \mathbb{R}^{256^2}$ to be the vector that represents the image and $y = Rx \in \mathbb{R}^{256}$ its Radon transformation. The DPC-CT measurements are now given by $b = Dy$. In order to generate this

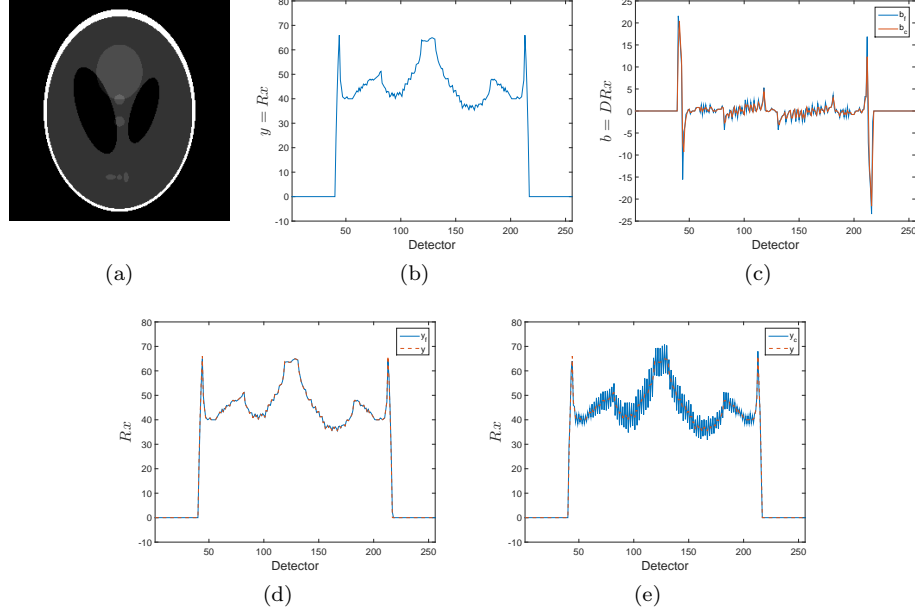


Figure 4: Data, measurements and reconstructions with model errors. (a) Shepp-Logan phantom. (b) Radon projection of x under a single projection angle ($\theta = \pi/2$). (c) The measurements b . (d) LSQR reconstruction using the forward difference model. (e) LSQR reconstruction using the central difference model.

data and avoid an inverse crime (see [25]), we generated a weighted version of b by taking

$$\begin{cases} b_f = (1 - \omega)D_f y + \omega D_c y \\ b_c = \omega D_f y + (1 - \omega)D_c y \end{cases} \quad (18)$$

for $\omega = 0.2$. For this dataset this yields a model error of approximately 10%. The data, the measurements and the reconstructions can be seen in Fig. 4. The reconstructions in Fig. 4d and Fig. 4e were made with 256 iterations of the LSQR algorithm. The results demonstrate that the central difference model is much more sensitive to the model errors we introduced by generating the data on a finer grid. Note that in this case both D_f and D_c are invertible, so the LSQR algorithm converged to $D_f^{-1}b$ and $D_c^{-1}b$ after the 256 iterations.

If, on the other hand, we assume that there are no model errors – i.e. calculate b without the weighting step – but add noise instead, similar results can be observed. This can be seen in Fig. 5 and Fig. 6, where we added 10% noise by generating a random normal vector $e \sim \mathcal{N}(0, 1)$ and setting

$$b = \tilde{b} + 0.10 \frac{\|b\|}{\|e\|} e, \quad (19)$$

for either the forward or the central difference model. Here, we also calculated the regularized solution with GBiT (using the maxcounter variable to perform 256 iterations).

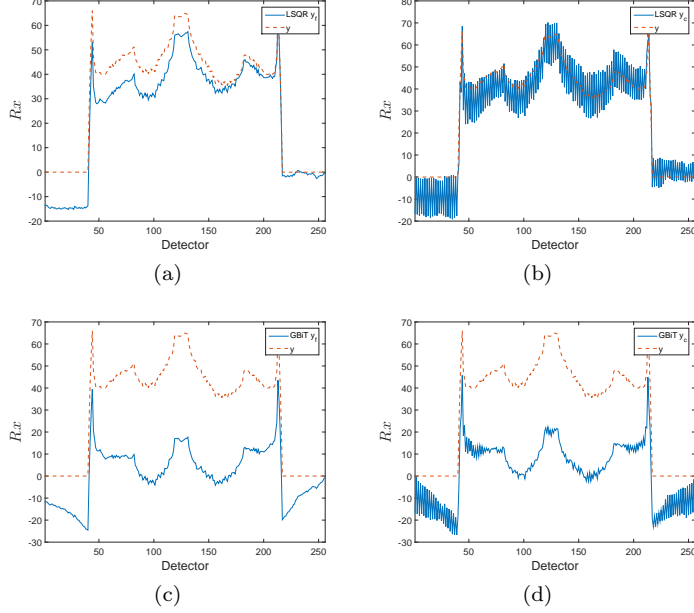


Figure 5: Reconstructions when noise is present. (a) LSQR reconstruction using the forward difference model. (b) LSQR reconstruction using the central difference model. (c) GBiT reconstruction using the forward difference model. (d) GBiT reconstruction using the central difference model.

When we compare the different solutions, we see that the central difference model is, again, more sensitive to the errors in the measurements. Furthermore, as can be seen in Fig. 5c and Fig. 5d, regularization hardly improves the oscillatory nature of the solution of the central difference model. Also, because we assumed no prior knowledge of the solution and used $x_0 = 0$ and $L = I$, the Tikhonov solution given by (9) will be centred around x_0 . This is because adding a constant function does not change the derivative and although these functions are not technically in the null space of either matrix, see proposition 1, they only fail at one or both of the boundary points. This will have little to no weight in the regularization.

Why the forward difference model gives much better reconstructions is explained by the inverse of the operators. If we write $b = \tilde{b} + e$ for some error caused by either the model, the measurements or both, the solution of $D_f y = b$ is given by a simple back substitution:

$$y_i = -\sum_{j=i}^n b_j = -\sum_{j=i}^n \tilde{b}_j - \sum_{j=i}^n e_j \quad (\forall i = 1, \dots, m). \quad (20)$$

So the noise partially cancels itself out. This is seen in Fig. 5a where the difference between the exact solution y and the approximation y_f does not increase monotonically from right to left. Important to note here is the assumption that the noise is centred around the measurements, i.e. $\mathbb{E}(b) = \tilde{b}$ or $\mathbb{E}(e) = 0$ for whatever distribution is used to model the noise. If this is not the case, e.g. the noise has a constant offset w.r.t the measurements, this offset will add up in the back substitution.

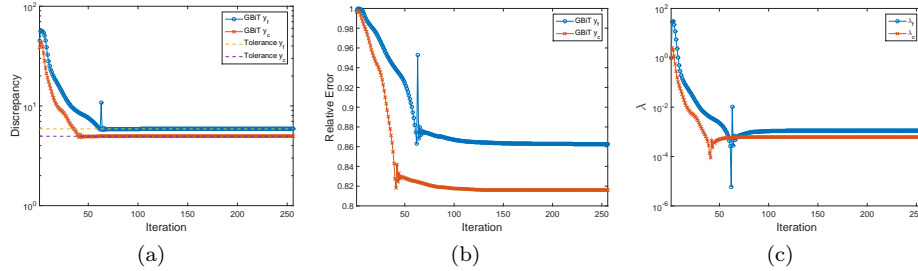


Figure 6: Overview of the discrepancy (a), relative error (b) and regularization parameter (c) in each iteration of GBiT for the reconstructions from Fig. 5c and Fig. 5d. For the forward and the central difference model the discrepancy principle is satisfied after respectively 80 and 43 iterations. The oscillations in all three plots at these iterations motivates performing some extra iterations after the discrepancy principle is satisfied for the first time.

This is illustrated in Fig. 7, which is the same problem from Fig. 5, but where the noise was given a constant offset of 5. Here, we observe that the regularization suppresses this offset because – as stated before – it tries to centre the solution around $x_0 = 0$. From the reconstructions, however, we see that the central difference model does not have the same smoothing effect on the solution. This is because when D_c is invertible (as is the case here), its inverse has a completely different structure¹:

$$D_c^{-1} = I_l \otimes \underbrace{\begin{pmatrix} 0 & -2 & 0 & -2 & 0 & -2 & \cdots \\ 2 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & -2 & 0 & -2 & \cdots \\ 2 & 0 & 2 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 0 & -2 & \cdots \\ 2 & 0 & 2 & 0 & 2 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}}_{k \times k} \in \mathbb{R}^{m \times m}$$

If we write down the analogous formulation to (20), a first observation is that, because the matrix consists of 0 and 2's, the noise components is multiplied by 2 as well. This implies that the part of the error that is not cancelled out doubles. Furthermore, since the rows have alternating plus and minus signs and a highly varying number of non-zero entries at alternating positions, the noise cancellation is less consistent compared to the forward difference model. This causes the heavy oscillations present in the non-regularized solutions.

5 Application to phase contrast tomography

The previous section looked at a single projection angle and only at the effect of the matrix D_f and D_c . We will now look at the full tomography problem for multiple

¹If L is a strictly lower bidiagonal matrix consisting of alternating zeros and twos, this matrix can be seen as the sum of L and $-L^T$.

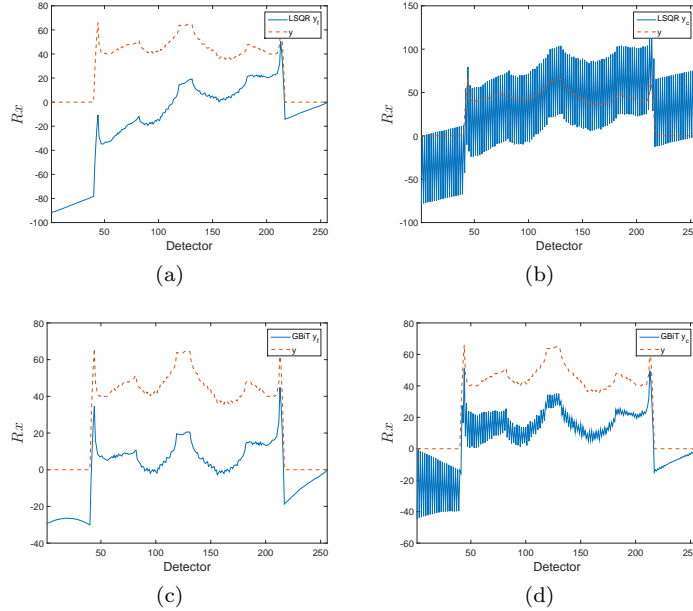


Figure 7: Reconstructions identical to Fig. 5, but with a constant offset added to the noise, i.e. e is replaced by $e + 5$ in (19).

angles in $[0, \pi]$. We start with a test problem and end with an application to real phase contrast data.

5.1 PC-CT with simulated data

We continue using the 256×256 Shepp-logan phantom from section 4.1 as a test image and measurement data is generated from 360 projection angles with 10% model error using the same weighting method as in (18) and add 10% noise to the resulting measurements as in (19). Because the inverse of D_f is given by a back substitution, we will also consider the reconstruction given by

$$Rx = D_f^{-1}b, \quad (21)$$

with and without regularization. This can be seen as two steps. First, reconstructing the Radon projections of δ for each projection angle from the measured phase contrast data and, second, approximate the image x based on the models for classic absorption based CT. This reconstruction is shown in Fig. 8a. However, since the classic tomography problem is known to be very ill-conditioned, this reconstruction is of low quality. Regularization, see Fig. 8d, suppresses this but over-smooths the reconstruction compared to directly solving $DRx = b$. This is demonstrated by the value of the regularization parameter λ , shown in Fig. 9b, which is significantly higher compared to the other values.

Fig. 8b and Fig. 8c demonstrate the results from the previous section, namely that the forward model is much less sensitive to the introduced measurement and model errors. This can be observed in Fig. 9a, which shows the relative error off all the reconstructions. Here it can be seen that the LSQR solution using the central

model shows signs of semi-convergence after 13 iterations, whereas for the solution using the forward model this only happens after 44 iterations and at a slower rate. When we compute the reconstruction using regularization, Fig. 8e and Fig. 8f, there is no semi-convergence is, but the reconstruction with the central model has less contrast due to the fact that it is more oscillatory than the other regularized reconstructions. This can be seen in the horizontal slices of the images, see Fig. 8g, Fig. 8h and Fig. 8i.

5.2 PC-CT with experimental data

Finally, to illustrate the validity of our test problem and to illustrate that GBiT is a viable algorithm for real data sets, we apply the same calculations on a PC-CT data set from experimental measurements provided by Franz Pfeiffer [31]. The data was taken over 2401 angles in $[0, \pi[$, with projections onto 1637 detectors. Because we have no prior knowledge about the noise, we use the alternative update scheme described in section 3.3. As can be seen from Fig. 10, Fig. 11 and Fig. 12, the results for the real data set are similar to our test problem. Here, all reconstructions were made by calculating 20 iterations of the algorithms, but for the GBiT algorithm applied to the forward model, the central model and (21), the discrepancy principle was satisfied after respectively 8, 9 and 7 iterations.

A final note should be added here about the runtime of the algorithm. The ASTRA toolbox can be used as a black box for the multiplication with R and R^T in the algorithms. This can be done on either the CPU or GPU. These reconstructions were done using the CPU code, with an average runtime of 3 minutes per iterations for the GBiT method. Comparing this to the runtime filtered back projection (20 seconds) this is a very poor result. However, when using the GPU the runtime of our iterative reconstruction drops significantly and the full algebraic reconstruction can be done in approximately 60 seconds. The reason we did not do this is because the ‘line weights’ we described in section 2.2 aren’t (yet) available for the GPU code. On the GPU only a ‘linear weighting’ scheme has been implemented that takes into account the neighbouring pixels. Furthermore, the GPU implementation of this scheme uses some approximations in the multiplication with R^T , which can make the reconstruction differ from the CPU reconstruction with the same weights.

6 Conclusions and Remarks

In summary, we adapted the generalized Arnoldi-Tikhonov method so that it can be applied to the general Tikhonov problem for non-square matrices. This was done by using the bidiagonal decomposition in order to rewrite the corresponding linear system. We tested our method by applying it to the PC-CT problem, where we demonstrated that the forward difference approximation resulted in much better reconstructions than the central difference approximation. Combining our results we were able to find high-contrast reconstructions coming from experimental data using the ASTRA toolbox. Although our CPU implementation is significantly slower than the FBP algorithm, we can improve the reconstruction time significantly using a GPU implementation.

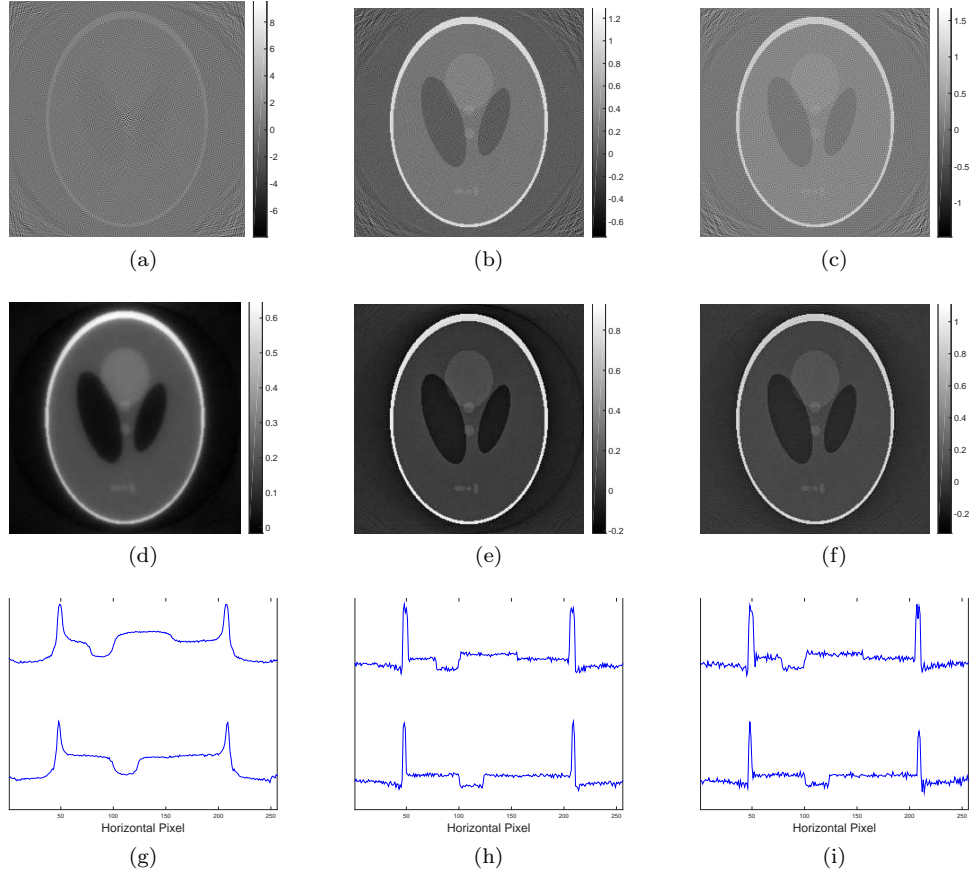


Figure 8: (a), (d) Reconstruction based on the phase retrieval approach from (21) without and with regularization. (b), (e) Reconstruction of $D_f Rx = b$ without and with regularization. (c), (f) Reconstruction of $D_c Rx = b$ without and with regularization. Note the different scales of the reconstructions. (g), (h) & (i) Normalized horizontal slices of the regularized reconstructions above them at vertical pixels 85 and 171 (top to bottom).

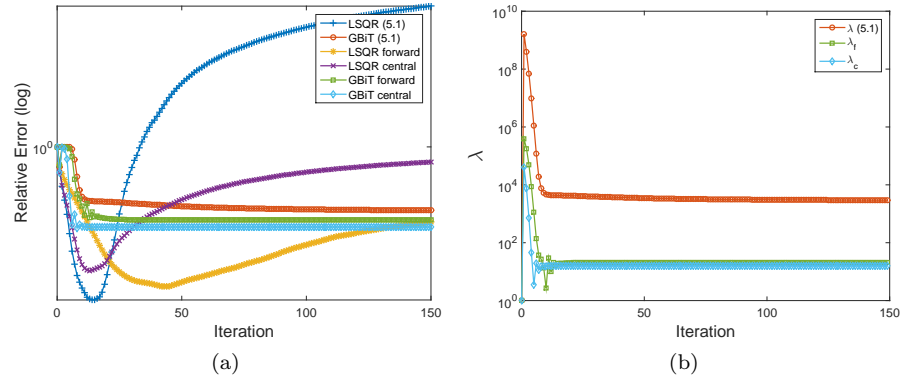


Figure 9: Relative error (a) and regularization parameter (b) for the reconstructions from Fig. 8.

Acknowledgements

We acknowledge Franz Pfeiffer, TUM, for providing the experimental data. NS also thanks the Department of Mathematics and Computer Science, U. Antwerpen for financial support. We also acknowledge fruitful discussions with Siegfried Cools.

References

- [1] A. Björck. *Numerical methods for least squares problems*. Siam, 1996.
- [2] U. Bonse and M. Hart. An x-ray interferometer with long separated interfering beam paths. *Applied Physics Letters*, 7(4):99–100, 1965.
- [3] M. Born and E. Wolf. *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light*. Cambridge university press, 1999.
- [4] A. V. Bronnikov. Theory of quantitative phase-contrast computed tomography. *JOSA A*, 19(3):472–480, 2002.
- [5] D. Calvetti, G. H. Golub, and L. Reichel. Estimation of the l-curve via lanczos bidiagonalization. *BIT Numerical Mathematics*, 39(4):603–619, 1999.
- [6] D. Calvetti, S. Morigi, L. Reichel, and F. Sgallari. Tikhonov regularization and the l-curve for large discrete ill-posed problems. *Journal of Computational and Applied Mathematics*, 123(1–2):423–446, 2000.
- [7] D. Calvetti, L. Reichel, and A. Shuibi. L-curve and curvature bounds for tikhonov regularization. *Numerical Algorithms*, 35(2–4):301–314, 2004.
- [8] J. Chung, J. G. Nagy, and D. P. O’Leary. A weighted gcv method for lanczos hybrid regularization. *Electronic Transactions on Numerical Analysis*, 28:149–167, 2008.

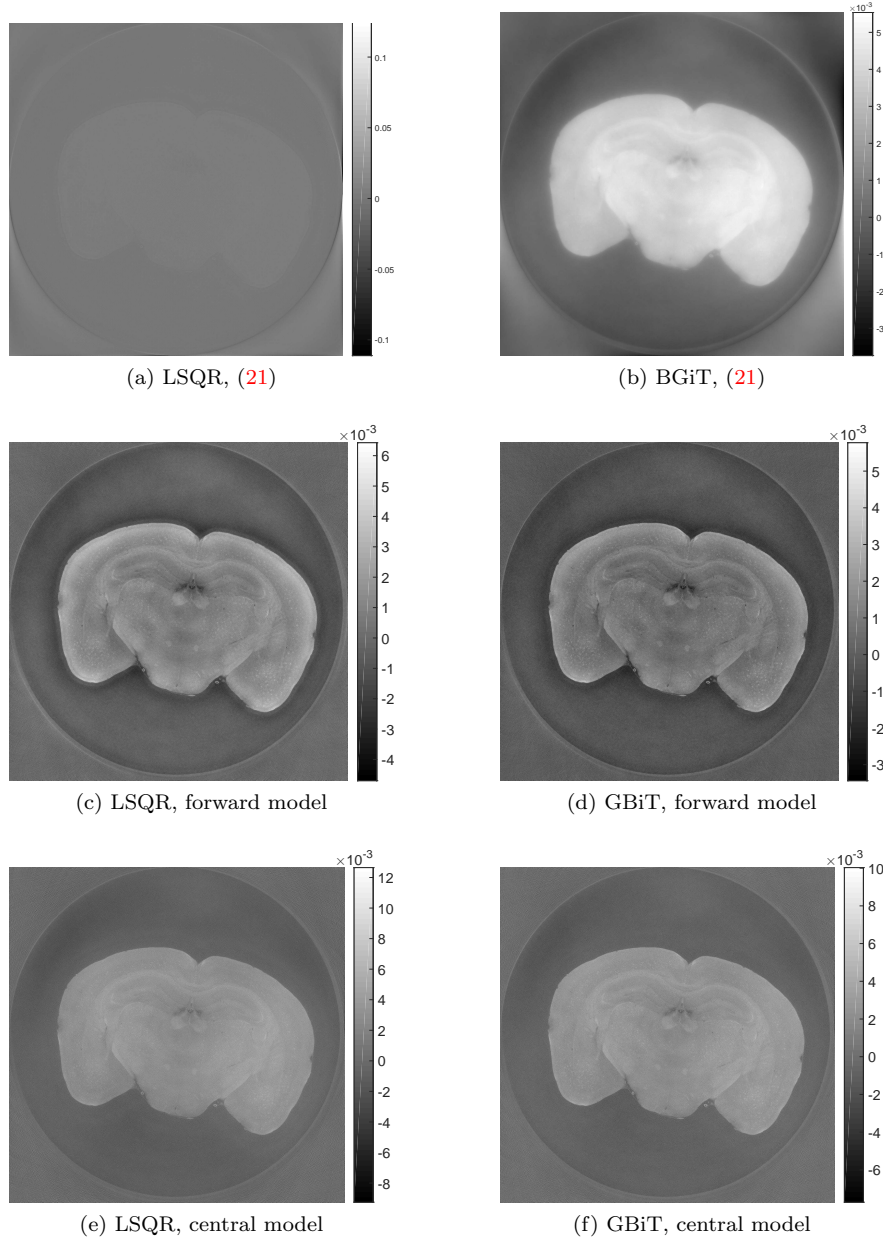
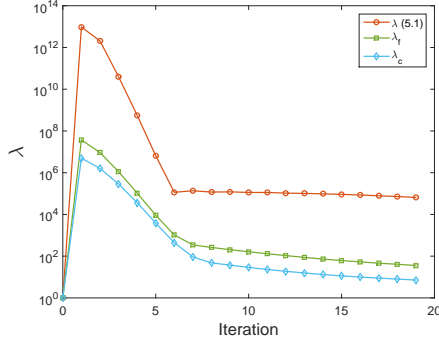
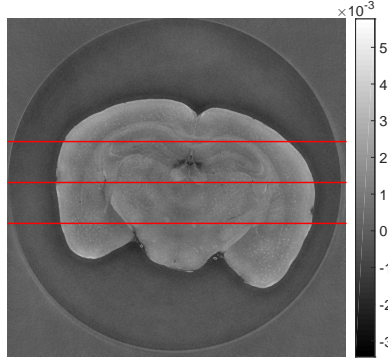


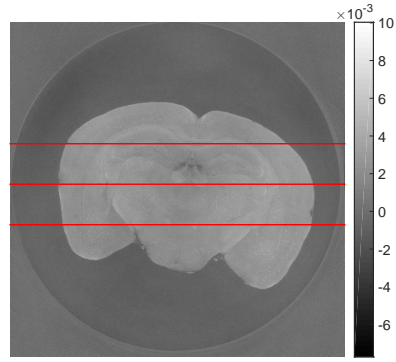
Figure 10: (a)(b) Reconstruction based on the phase retrieval approach from (21) without and with regularization. (c), (d) Reconstruction of $D_f Rx = b_n$ without and with regularization. (e), (f) Reconstruction of $D_c Rx = b_n$ without and with regularization.



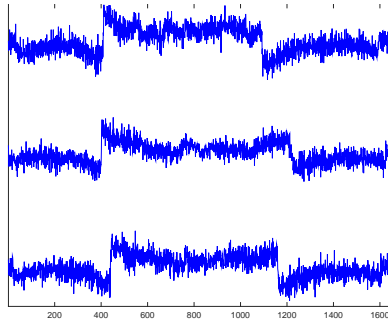
(a)



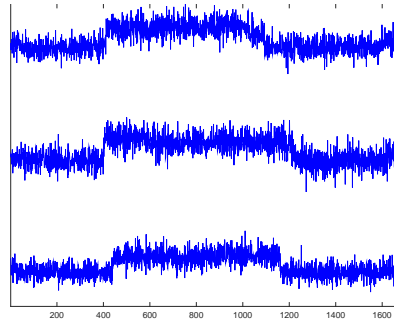
(b) GBiT, forward model



(c) GBiT, central model



(d)



(e)

Figure 11: (a) Regularization parameter λ . (b)-(e) Normalized horizontal slices of the regularized reconstructions above them at the denoted vertical positions. The higher oscillations for the the reconstruction based on the central model result in a lower contrast.

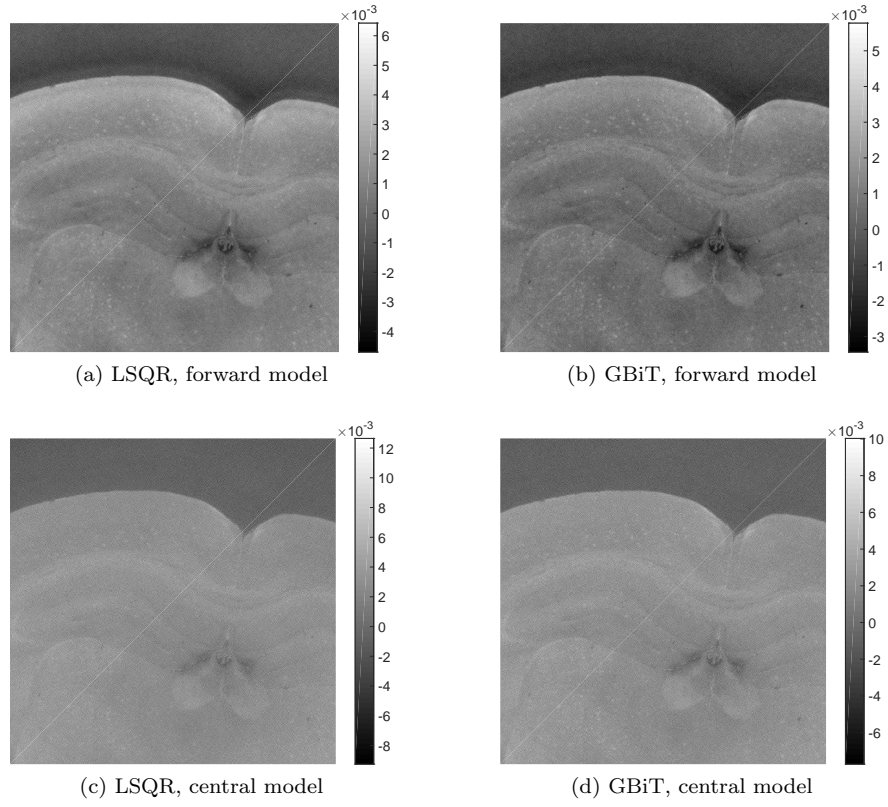


Figure 12: Details from the various reconstructions: (a) detail from 10c, (b) detail from 10d, (c) detail from 10e, (d) detail from 10f.

- [9] J. Fu, X. Hu, A. Velroyen, M. Bech, M. Jiang, and F. Pfeiffer. 3d algebraic iterative reconstruction for cone-beam x-ray differential phase-contrast computed tomography. *PloS one*, 10(3):e0117502, 2015.
- [10] J. Fu, S. Schleede, R. Tan, L. Chen, M. Bech, K. Achterhol, M. Gifford, R. Loewen, R. Ruth, and F. Pfeiffer. An algebraic iterative reconstruction technique for differential x-ray phase-contrast computed tomography. *Zeitschrift für Medizinische Physik*, 23(3):186–193, 2013.
- [11] J. Fu, R. Tan, and L. Chen. Analysis and accurate reconstruction of incomplete data in x-ray differential phase-contrast computed tomography. *Analytical and bioanalytical chemistry*, 406(3):897–904, 2014.
- [12] S. Gazzola and J. G. Nagy. Generalized arnoldi-tikhonov method for sparse reconstruction. *SIAM Journal on Scientific Computing*, 36(2):B225–B247, 2014.
- [13] S. Gazzola and P. Novati. Automatic parameter setting for arnoldi-tikhonov methods. *Journal of Computational and Applied Mathematics*, 256:180–195, 2014.
- [14] S. Gazzola, P. Novati, and M. R. Russo. On krylov projection methods and tikhonov regularization. *Electron. Trans. Numer. Anal. To appear*, 2014.
- [15] G. H. Golub, M. Heath, and G. Wahba. Generalized cross-validation as a method for choosing a good ridge parameter. *Technometrics*, 21(2):215–223, 1979.
- [16] G. H. Golub and W. Kahan. Calculating the singular values and pseudo-inverse of a matrix. *Journal of the Society for Industrial & Applied Mathematics, Series B: Numerical Analysis*, 2(2):205–224, 1965.
- [17] P. C. Hansen. *Discrete inverse problems: insight and algorithms*, volume 7. Siam, 2010.
- [18] V. N. Ingal and E. A. Beliaevskaya. X-ray plane-wave topography observation of the phase contrast from a non-crystalline object. *Journal of Physics D: Applied Physics*, 28(11):2314, 1995.
- [19] T. J. Davis, D. Gao, T. E. Gureyev, A. W. Stevenson, and S. W. Wilkins. Phase-contrast imaging of weakly absorbing materials using hard x-rays. *Nature*, 373(6515):595–598, 1995.
- [20] P. M. Joseph. An improved algorithm for reprojecting rays through pixel images. *IEEE Transactions on Medical Imaging*, 1(3):192–196, 1982.
- [21] A. Kostenko, K. J. Batenburg, A. King, S. E. Offerman, and L. J. van Vliet. Total variation minimization approach in in-line x-ray phase-contrast tomography. *Optics express*, 21(10):12185–12196, 2013.
- [22] A. Momose, S. Kawamoto, I. Koyama, Y. Hamaishi, K. Takai, and Y. Suzuki. Demonstration of x-ray talbot interferometry. *Japanese journal of applied physics*, 42(7B):L866, 2003.

- [23] A. Momose, T. Takeda, and Y. Itai. Phase-contrast x-ray computed tomography for observing biological specimens and organic materials. *Review of scientific instruments*, 66(2):1434–1436, 1995.
- [24] A. Momose, W. Yashiro, Y. Takeda, Y. Suzuki, and T. Hattori. Phase tomography by x-ray talbot interferometry for biological imaging. *Japanese journal of applied physics*, 45(6R):5254, 2006.
- [25] J. L. Mueller and S. Siltanen. *Linear and Nonlinear Inverse Problems with Practical Applications*. SIAM, 2012.
- [26] F. Natterer et al. *Mathematical methods in image reconstruction*. Siam, 2001.
- [27] M. Nilchian, C. Vonesch, P. Modregger, M. Stampanoni, and M. Unser. Fast iterative reconstruction of differential phase contrast x-ray tomograms. *Optics express*, 21(5):5511–5528, 2013.
- [28] C. C. Paige and M. A. Saunders. Algorithm 583: LSQR: Sparse linear equations and least squares problems. *ACM Transactions on Mathematical Software*, 8(2):195–209, 1982.
- [29] C. C. Paige and M. A. Saunders. LSQR: An algorithm for sparse linear equations and sparse least squares. *AMC Transactions On Mathematical Software*, 8(2):43–71, 1982.
- [30] W. J. Palenstijn, K. J. Batenburg, and J. Sijbers. Performance improvements for iterative electron tomography reconstruction using graphics processing units (GPUs). *Journal of Structural Biology*, 176:250–253, 2011. <http://sourceforge.net/p/astra-toolbox/wiki/Home/>.
- [31] F. Pfeiffer, O. Bunk, C. David, M. Bech, G. L. Duc, A. Bravin, and P. Cloetens. High-resolution brain tumor visualization using three-dimensional x-ray phase contrast tomography. *Physics in medicine and biology*, 52(23):6923, 2007.
- [32] F. Pfeiffer, T. Weitkamp, O. Bunk, and C. David. Phase retrieval and differential phase-contrast imaging with low-brilliance x-ray sources. *Nature physics*, 2(4):258–261, 2006.
- [33] A. Snigirev, I. Snigireva, V. Kohn, S. Kuznetsov, and I. Schelokov. On the possibilities of x-ray phase contrast microimaging by coherent high-energy synchrotron radiation. *Review of Scientific Instruments*, 66(12):5486–5492, 1995.
- [34] M. Takeda, H. Ina, and S. Kobayashi. Fourier-transform method of fringe-pattern analysis for computer-based topography and interferometry. *JosA*, 72(1):156–160, 1982.
- [35] W. van Aarle, W. J. Palenstijn, J. D. Beenhouwer, T. Altantzis, S. Bals, K. J. Batenburg, and J. Sijbers. The astra toolbox: A platform for advanced algorithm development in electron tomography. *Ultramicroscopy*, 2015.
- [36] P. Zhu, K. Zhang, Z. Wang, Y. Liu, X. Liu, Z. Wu, S. A. McDonald, F. Marone, and M. Stampanoni. Low-dose, simple, and fast grating-based x-ray phase-contrast imaging. *Proceedings of the National Academy of Sciences*, 107(31):13576–13581, 2010.